



PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification 6 :</b> <b>C07H 21/02, 21/04, C12N 15/00</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 95/15972</b> <b>(43) International Publication Date:</b> 15 June 1995 (15.06.95)
<b>(21) International Application Number:</b> PCT/US94/14181 <b>(22) International Filing Date:</b> 9 December 1994 (09.12.94) <b>(30) Priority Data:</b> 164,303 9 December 1993 (09.12.93) US <b>(71) Applicant:</b> THOMAS JEFFERSON UNIVERSITY [US/US]; 1020 Locust Street, Philadelphia, PA 19107 (US). <b>(72) Inventor:</b> KMIEC, Eric, B.; 1 Lapp Road, Malvern, PA 19355 (US). <b>(74) Agents:</b> FRIEBEL, Thomas, F. et al.; Pennie & Edmonds, 1155 Avenue of the Americas, New York, NY 10036 (US).		<b>(81) Designated States:</b> AM, AU, BB, BG, BR, BY, CA, CN, CZ, EE, FI, GE, HU, JP, KE, KG, KR, KZ, LK, LR, LT, LV, MD, MG, MN, MW, NO, NZ, PL, RO, RU, SD, SI, SK, TJ, TT, UA, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ).  <b>Published</b> <i>With international search report.</i>
<b>(54) Title:</b> COMPOUNDS AND METHODS FOR SITE-DIRECTED MUTATIONS IN EUKARYOTIC CELLS <b>(57) Abstract</b> <p>The present invention concerns a polynucleotide having both ribonucleotides and deoxyribonucleotides in a first strand and solely deoxyribonucleotides in a second strand; wherein the strands are Watson-Crick paired and are linked by an oligonucleotide so that the polynucleotide has at most a single 3' and a single 5' end. These ends can be ligated so that the polynucleotide is a single continuous circular polymer. The polynucleotide can be used to induce specific alterations in targeted genes.</p>		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo			SE	Sweden
CH	Switzerland	KR	Republic of Korea	SI	Slovenia
CI	Côte d'Ivoire	KZ	Kazakhstan	SK	Slovakia
CM	Cameroon	LI	Liechtenstein	SN	Senegal
CN	China	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TG	Togo
CZ	Czech Republic	LV	Latvia	TJ	Tajikistan
DE	Germany	MC	Monaco	TT	Trinidad and Tobago
DK	Denmark	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	US	United States of America
FI	Finland	ML	Mali	UZ	Uzbekistan
FR	France	MN	Mongolia	VN	Viet Nam
GA	Gabon				

## COMPOUNDS AND METHODS FOR SITE-DIRECTED MUTATIONS IN EUKARYOTIC CELLS

The present application is a Continuation-in-Part of  
5 patent application Serial No. 08/164,303, filed December 9,  
1993, which is hereby incorporated by reference in its  
entirety.

### 1. FIELD OF THE INVENTION

10 The present invention concerns the field of molecular  
genetics. Particularly it concerns nucleic acid compounds  
and methods of their use to introduce specific genetic  
alterations into living cultured eukaryotic cells. More  
particularly it concerns substantially Watson-Crick paired,  
15 duplex nucleic acids wherein a portion of one strand of the  
duplex comprises a 2'-O or 2'-OMe ribose containing  
nucleotides and the remainder comprise deoxyribose  
nucleotides.

20

### 2. BACKGROUND TO THE INVENTION

#### 2.1. CHIMERIC AND/OR HYBRID DUPLEX NUCLEIC ACIDS

The field of the invention concerns nucleic acids.  
25 Nucleic acids are heteropolymers, i.e., polymers of non-  
identical subunits, which are linked by oriented  
phosphodiester bonds or their derivatives, into polymers.  
Duplex nucleic acids are nucleic acids wherein each base of a  
first strand of the duplex corresponds to a base of a second  
30 strand of the duplex according to the scheme in which uracil  
or thymine and adenine correspond and cytosine and guanine  
correspond. Anti-parallel duplex strands having these  
correspondences are said to be Watson-Crick paired. Duplex  
nucleic acids can be of two major types, ribonucleic acids  
35 and deoxyribonucleic acids. Each ribonucleotide has an  
equivalent deoxyribonucleotide, e.g., adenosine and  
deoxyadenosine, cytidine and deoxycytidine, guanosine and

deoxyguanosine, uridine and thymidine. As used in the field, a nucleic acid in which both ribonucleotides and deoxyribonucleotides are present in the same strand is termed a mixed or chimeric (hereinafter "chimeric") nucleic acid. A  
5 duplex nucleic acid in which deoxyribonucleotides and ribonucleotides correspond with each other is termed a hybrid-duplex. When two strands form a region of duplex nucleic acid for less than all of their bases, the resultant molecule is termed a heteroduplex.

10 Most often, the two strands of a duplex nucleic acid are not covalently bonded but are associated only by Watson-Crick pairing. However, the two strands of a duplex can be linked by an oligonucleotide to form a single polymer. The linking oligonucleotide is not Watson-Crick paired. The heteroduplex  
15 in which the first and second strands are portions of a single polymer is termed a "hairpin duplex" or a "stem and loop" structure. The former term will be used hereinafter.

As used herein chimerism is a property of a nucleic acid polymer and hybridism is a property of a duplex. For  
20 ~~example, an mRNA and its template form a hybrid duplex though~~ neither is chimeric, while, for example, the chimeric octanucleotides 5'd(TTTT)-r(CCCC)3' and 5'r(GGGG)-d(AAAA)3' will form a Watson-Crick duplex with each other but the resultant duplex is not a hybrid-duplex. A duplex nucleic  
25 acid which is not a hybrid-duplex is termed hereinafter a "homo-duplex". Unless specifically noted otherwise, a homo-duplex nucleic acid refers only to a deoxynucleotide containing duplex. Lastly, note that while those in the field refer to the formation of a Watson-Crick duplex as  
30 "hybridization," even where there is no hybrid-duplex nucleic acids.

Those interested in the study the structure of chimeric and/or hybrid duplex nucleic acids by X-ray diffraction and 2-dimensional NMR have synthesized chimeric nucleic acids and  
35 hybrid duplex nucleic acids for use in their studies. See, e.g., Salazar, M., et al., 1994, J.Mol.Biol. 241:440-55 and Egli, M., et al., 1993, Biochemistry 32: 3221-37 (two

stranded chimeric hybrid duplex of the form  $r_3d_7 \bullet d_{10}$ ); Ban, C., et al., 1994, J.Mol.Biol. 236:275-85 (self complementary chimeric hybrid duplex of the form  $d_3r_1d_4$ ); Chou, S.H., 1991, Biochemistry 30:5248-57 (self-complementing and non-self-complementing chimeric hybrid duplexes of the form  $d_4r_4d_4$ ). The complementary strands of these duplex nucleic acids were not covalently bound to each other; they were associated only by Watson-Crick pairing.

A second group of scientists who have synthesized chimeric nucleic acids are those interested in the study and use of ribozymes, i.e., RNA molecules that are either self-cleaving or cleave other RNAs. Perreault, J.P., et al., 1990, Nature 344:565; Taylor, N.R., et al., 1992, Nucleic Acids Research 20:4559-65; Shimaya, T., 1993, Nucleic Acids Research 21:2605. These researches have found that chimeric ribozymes are active and are more resistant to nuclease digestion than RNA ribozymes. Chimeric ribozymes are self-complementary, i.e., the Watson-Crick paired strands are covalently linked. The compounds synthesized during the studies of chimeric ribozymes differ from the above-noted hybrid-duplex molecules, that were synthesized used for structural studies, in that chimeric ribozymes do not contain stable hybrid-duplexes. Thus, a chimeric ribozyme having DNA binding arms binds to its substrate and forms a hybrid duplex. Yang, J.H., et al., 1990, Biochemistry 29:11156-60. See also, Sawata, S., et al., 1993, Nucleic Acids Research 21:5656-60; Hendry, P., et al., 1992, Nucleic Acids Research 20:5737-41 Shimayama, T., 1993, Nucleic Acids Research 21:2605. The ribozyme catalyzes the cleavage of the RNA substrate and the hybrid-duplex is thus dissolved.

## 2.2. SITE-DIRECTED GENETIC ALTERATION IN EUKARYOTIC CELLS

Those skilled in the art of molecular biology recognize that on frequent occasions it is desired not merely to introduce a new polynucleic acid sequence, i.e., a new gene, into a targeted eukaryotic cell, but rather to alter a

defined, pre-existing gene in the targeted cell. The targeted cell can be used in culture or it can be used to construct a transgenic animal.

A wide variety of techniques have been developed to  
5 introduce DNA into cultured eukaryotic cells. These techniques include calcium phosphate precipitation and DEAE-dextran mediated endocytosis, electroporation, liposome mediated fusion and transduction with replication incompetent viruses. However, while such techniques can quite often  
10 introduce functional genes into the eukaryotic cell, these techniques do not readily accomplish an alteration (mutation) in a specific existing gene. After introduction the exogenous DNA is isolated at a random position in the cell's genome by illegitimate recombination, rather than at a specific  
15 position by homologous recombination.

To date there is no generally satisfactory scheme for introducing a site-directed or site-specific genetic alteration (mutagenesis) in a higher eukaryote, i.e., in mammalian or avian cells. Although homologous recombination  
20 can be obtained in higher eukaryotic cells by introduction of very long (> 1 kb) nucleic acids, these techniques require the application of elaborate selection techniques because the rate of illegitimate recombination in higher eukaryotes greatly exceeds that of homologous recombination. Thomas,  
25 K.R. & Capecchi, M.R., 1987, Cell 52:503. See, also, Valancius, V. & Smithies O., 1991, Mol. Cell. Biol. 11:4389 (comparison homologous recombination of linearized and supercoiled plasmids in eukaryotic cells).

One approach to achieving a predominantly site-directed  
30 mutagenesis has been the introduction of single stranded oligodeoxynucleotides directly into the cell. This techniques has been successfully employed in the yeast *Saccharomyces cerevisiae*, in which homologous recombination is significantly more active than it is in higher eukaryotes.  
35 Moerschell, R.P., et al., 1988, Proc.Natl.Acad.Sci. 85:524-28; Yamamoto, T., et al., 1992, Yeast 8:935-48. However, to date there have been no reports of the successful

transformation of mammalian or avian cells by single stranded oligonucleotides.

A relationship between the structure of the target DNA and the rate of homologous recombination in mammalian can be  
5 inferred by studies that show that regions of alternating purine and pyrimidine bases, i.e., [d(TG)<sub>30</sub>•d(AC)<sub>30</sub>], display an entranced rate of recombination. These effects were demonstrated in studies of non-replicating plasmids in cultured mammalian cells. Wahls, W.P., et al., 1990, Mol.  
10 Cell. Biol. 10:785-93. These experiments were not extended to show recombination between an exogenous nucleic acid and the genome of the cell.

Attempts have been made to use RecA, a protein that promotes homologous recombination in the bacteria, *E. coli*,  
15 to promote homologous recombination in eukaryotic cells. However, these attempts have not been clearly successful. For example U.S.Pat No. 4,950,599 to W. Bertling discloses a very low rate of site-directed mutation and no enhancement in the rate of homologous recombination by use of RecA in  
20 eukaryotic cells. Patent publications WO 93/22443 to D. Zarling and E. Sena, and publication 94/04032 to D.C. Gruenert and K. Kunzelmann both purport to correct a genetic defect in a cultured cell line related to cystic fibrosis. These publications disclose primarily experimental data that  
25 demonstrate the principle rather than data concerning examples of operative methods. Thus, to introduce polynucleotide/RecA complexes access to the nucleus, Zarling and Gruenert employ cells that were membrane-permeabilized, although such cells are incapable of further growth.  
30 Moreover, even when RecA-promoted homologous recombination was asserted to have taken place in intact cells, these publications provide no quantitative estimates of its frequency. Thus, the use of prokaryotic recA has not been convincingly shown to result in a rate homologous  
35 recombination in any viable eukaryotic cell significantly greater than the spontaneous rate of homologous recombination.

## 3. SUMMARY OF THE INVENTION

The present invention concerns single covalently linked duplex oligonucleotides that are homologous with a gene of interest having both deoxyribonucleotides and ribonucleotides. In order to effect a genetic change there are within the region of homology one or more non-corresponding (hereinafter "heterologous" or "mutator") base pairs. The normal, constitutive cellular processes of homologous recombinants cause the mutator nucleotides to be inserted into the targeted genomic site. The duplex oligonucleotides of the invention (hereinafter "chimeric vectors") can be used to alter specifically a gene of interest by introducing into the gene the heterologous base pairs. The heterologous base pairs can be base pairs different from the gene of interest, or base pairs in addition to those present in the gene of interest (an insertion), or, lastly, the heterologous base pairs can be the absence of base-pairs found in the gene of interest (a deletion). The present invention is based, in part, on the discovery that the inclusion of a region of between about 15 and 50 base pairs of hybrid-duplex nucleic acid in the vector causes a greatly increased rate of alteration of the gene of interest. When the region of the heterologous base pairs is between 1 and 50 base pairs, the heterologous base pairs can be present in the vectors of the invention as either a homo- or a hybrid-duplex. When the heterologous base pairs are greater than 50 base pairs in length it is preferred that they be present as a homo-duplex.

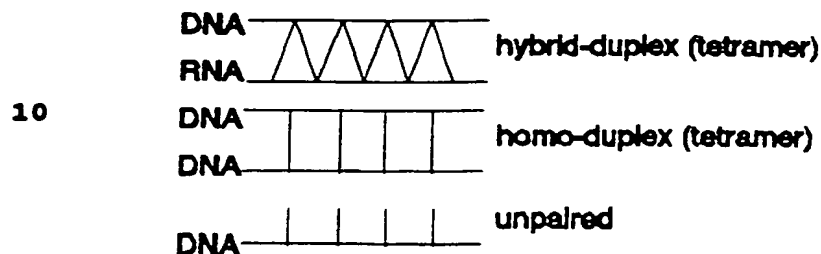
The vector can be introduced into the target cell by any method known to allow for the introduction of nucleic acids into eukaryotic cells. Without limitation as to theory, the chimeric vector is believed to be engaged by the recombination/repair mechanisms of the target cell and to direct a the alteration of the target gene by gene conversion or by homologous recombination.



#### 4. BRIEF DESCRIPTION OF THE FIGURES

Figure 1 shows the schematic representation of two chimeric vectors. The following are particularly labeled: 1, a first strand; 2, a second strand; 3, a heterologous region; 4, a homologous region; 5, a linker; 6, a ligation site.

##### Symbol Key for Figure 1



15

Figure 1A is a ligated chimeric vector in the R-D-R form.

20

Figure 1B is a hairpin chimeric vector in the R-D-R form having a single 3' and 5' end.

#### 5. DETAILED DESCRIPTION OF THE INVENTION

25

##### 5.1. DESCRIPTION OF THE CHIMERIC VECTORS OF THE INVENTION

30 The present invention encompasses duplex nucleic acids that contain both deoxyribose and ribose containing bases. Thus they contain regions of both DNA and RNA and are termed therefore termed "chimeric vectors." The 2'-O of the ribonucleotides of the chimeric vector can be methylated. Any phosphodiester can be substituted by a phosphorothiodiester or methyphosphonodiester.

35 The chimeric vectors of the invention are a single nucleic acid polymer. Accordingly, the chimeric vectors of the invention must contain at least one region of between at

least 1 base and more usually 3 or 4 bases that are not Watson-Crick paired. These unpaired regions serve as linkers between the two strands of Watson-Crick paired bases. In contrast to other chimeric nucleotides that have been  
5 synthesized having regions of unpaired nucleotides chimeric vectors have no enzymatic activity i.e., they do not catalyze chemical reactions or themselves undergo chemical reaction in the absence of an biological energy source such as ATP.

The bases of the linkers in the preferred embodiment are  
10 deoxyribonucleotides. A chimeric vector having two linkers can be constructed so that the 3' and 5' ends of the polymer are Watson-Crick paired to adjacent nucleotides of the complementary strand. They can then be readily ligated so that the chimeric vector forms a single continuous circular  
15 nucleic acid polymer.

Substantially all the remaining bases of the chimeric vector are Watson-Crick paired. As used herein the statement that bases are Watson-Crick paired or that they form a duplex nucleic acid is to be understood to mean that  
20 ~~under the proper conditions of temperature and salt they are~~ capable of forming base pairs or a duplex nucleic acid. It is to be understood that under some conditions of low salt and/or elevated temperature the Watson-Crick base pairs may cease to be thermodynamically stable such that the duplex  
25 nucleic acid melts-out or denatures. It is also to be understood that a one or two base pair mismatch does not effect the operability of the invention.

The chimeric vectors of the present invention are intended for the purpose of specifically introducing  
30 alterations (a mutation) into a target gene. The genetic site of the alteration is determined by selecting a portion of the chimeric vector to have the same sequence as (to be homologous with) the sequence of the target site, hereinafter termed a homologous region. The area of differences between  
35 the sequence of the chimeric vector and the target gene are termed the heterologous region. Note that the chimeric vector is heterologous to the target gene, but is not a

heteroduplex in this region of the vector. According to the preferred embodiment of the invention there are two homologous regions in each chimeric vector flanking the interposed heterologous region, all three regions being  
5 present in a single continuous duplex nucleic acid. Further according to the invention each homologous region contains a portion of hybrid-duplex nucleic acid. The portion of each hybrid-duplex can be at least 4 base pairs and preferably is 8 base pairs and more preferably about 20 to 30 base pairs.  
10 The function of the chimeric vector is not effected by a dinucleotide base pair of homo-duplex within a region of hybrid duplex, placed to allow ligation of the 3' and 5' ends to each other. The total length of the two homologous regions can be at least 20 base pairs and preferably is  
15 between 40 and 60 base pairs.

According to the invention, a region of homo-duplex nucleic acid can be disposed between the hybrid-duplex / homologous regions of the vector. The interposed homo-duplex can contain the heterologous region. When the heterologous  
20 region is less than about 50 base pairs and preferably less than about 20 base pairs, the presence of a interposed homo-duplex region is optional. When the heterologous region exceeds about 20 base pairs, an interposed homo-duplex that contains the heterologous region is preferred.

25

## 5.2. CONSTRUCTION OF THE CHIMERIC VECTORS OF THE INVENTION

The chimeric vectors of the invention can be synthesized by any method. The chimeric vectors can most readily be  
30 synthesized by modification of the techniques used in the solid phase synthesis of DNA. Reviewed Caruthers, M.H., 1985, Science 230:281-85; Itakura, K., et al., 1984, Ann.Rev.Biochem. 53:523-56. Modifications to permit the synthesis of RNA and of chimeric nucleic acids are disclosed  
35 in Scaringe, S.A., et al., 1990, Nucleic Acids Research 18:5433-41; Usman, N., et al., 1992, Nucleic Acids Research 20:6695-99; and Swiderski, P.M., et al., 1994, Anal. Biochem

216:83-88, which are hereby incorporated by reference in their entirety. Recent advances concerning the synthesis of chimeric nucleic acids are reviewed in Usman, N. & Cedergren, R., 1992, Trends Bioch. Sci. 17:334-9.

5 Chimeric vectors having a homo-duplex region interposed between two hybrid-duplex regions can be constructed using semisynthetic techniques. Two synthetic chimeric polynucleic acids having a hairpin conformation are to be constructed. The free 5' and 3' ends of the two chimeric nucleic acids are  
10 constructed with an overlap staggered ends complementary to the overlap of two different restriction enzyme digest products. A homo-duplex region is provided having the complementary restriction enzyme digested ends. The addition of a restriction enzyme sites to the ends of a cloned DNA  
15 fragment can be accomplished by techniques well understood by those skilled in the art, e.g., without limitation, PCR amplification with extended primers or the blunt end ligation of linkers containing the restriction site. The two chimeric nucleotides and the homo-duplex region can be ligated by  
20 ~~conventional enzymatic techniques.~~ The product, having chimeric oligonucleotides ligated at both ends can be separated from the incompletely reacted substrates by electrophoresis in 6% polyacrylamide gel in Tris Borate EDTA buffer under non-denaturing conditions. The linear capped  
25 molecules are constrained and are electrophoresed more slowly under these conditions.

Chimeric vectors containing only naturally occurring nucleotides can be used for the present invention. The chimeric vectors having regions of hybrid-duplex of about 20  
30 ribonucleotides are found, *in vitro*, to be resistant to RNase H. Resistance to enzymatic degradation can be obtained by the replacement of the ribonucleotides with 2'-O methylated ribonucleotides. Additionally or alternatively the phosphodiester bonds of the nucleic acid can be replaced by  
35 phosphorothiodiesters. Shimayama, T. et al., 1993, Nucleic Acids Research 21:2605. Arabinose containing nucleotides can also be used. As used herein the term nucleic acid is

intended to encompass nucleic acids having these modifications.

### 5.3. THE USE OF THE CHIMERIC VECTORS OF THE INVENTION

5

The chimeric vectors of the present invention can be used to a mutation in a specific location in the genome of a target cell. The specific location of the target location is defined by its nucleic sequence hereinafter the target  
10 sequence. According to the invention a chimeric vector is constructed wherein the homology regions are identical to the target sequence, except for the presence of some regions of hybrid-duplex. The change to be introduced is encoded by the heterologous region. The change may be change in one or more  
15 bases of the sequence or the addition of one or more bases. When the change in the target sequence is the addition of less than about 20 bases the invention may be practiced using one or two regions of hybrid duplex. When the change in the target sequence is the addition of more than about 50 bases  
20 it is preferred that the heterologous region be contained within a homo-duplex region.

The practice of the invention requires that the chimeric vector be introduced into the nucleus of the target cell. Any method which causes such introduction can be used. Such  
25 methods include electroporation, DEAE-dextran,  $\text{Ca PO}_4$  precipitation, liposome mediated fusion (LIPOFECTIN), and direct injection. Electroporation is particularly suitable.

In one embodiment of the invention the chimeric vector can be used to construct transgenic animals. The chimeric  
30 vector can be introduced into the pronucleus of a ovum by direct injection, according to the method described Brinster, R.L. et al., 1989, PROC.NATL.ACAD.SCI 86:7087; see also U.S. Patent No. 4,873,191 to T.E. Wagner and P.C. Hoppe, which are hereby incorporated by reference in their entirety.

35 Alternatively, the chimeric vector can be introduced into an embryonic stem cell, chimeric animals can be produced by aggregation of the embryonic stem cell with normal blastocyst

cells, and transgenic animals can be recovered as offspring of the chimeric animals, according to the method of Capecchi, M.R., 1989, Science 244: 1288, which is hereby incorporated by reference in its entirety.

5        Using electroporation, as many as 1 cell per 10,000 treated cells can be specifically mutated at the target sequence (hereinafter "transformed"). The practice of the invention, thus, includes the use of a method to select the transformed cells from among the larger number of unmutated  
10 cells. In one embodiment the transformation of the cells confers a growth advantage. Non-limiting examples of such growth advantages include drug-resistance, alterations in growth regulation, and alterations in the capacity to utilize metabolites. In an alternative embodiment the method of  
15 selection can be negative selection whereby the transformed cells are rendered incapable of growth under the selecting conditions and the non-transformed cells are removed by exposure to conditions which selectively destroy proliferating cells.

~~20        Alternatively, the transformed cells may have an altered cell-surface antigenic phenotype that can be detected by immunofluorescence and selection can be performed by a Fluorescence Activated Cell Sorter.~~

When the method of introducing the chimeric vector into  
25 the cell is direct injection, as for example when constructing transgenic animals by pronuclear injection, the rate of transformation can be greater than 1 per 10,000 cells. The need for selection is thereby considerably reduced.

30        Typically useful amounts of a chimeric vector are between 10 and 1000 ng of chimeric vector per million cultured cells to be transformed by electroporation.

## 6. EXAMPLES

6.1. **Example 1: In Vivo Activity in *Ustilago Fungus***

Wild-type *Ustilago* has a functioning *ura-3* gene whose product is part of the uracil biosynthetic pathway. When wild-type *Ustilago* is grown on 5'-fluororotic acid (5FOA)-media the cells die due to incorporation of the acid into the pathway. If the *ura-3* gene is mutated so that *ura-3* mRNA is not produced, the cells survive on the 5FOA media.

The sequence of the endogenous *ura-3* gene is known in the art. In one set of experiments, base 358 of the sequence was changed from a thymidine to an adenine which mutation results in a dysfunctional protein.

A chimeric vector was synthesized as follows:

Vector for mutation at base 358: (358 RNA Vector)

5' TGCCGATCGGCAACTTTTGUUGCCGAUCGGCAAATTT 3' (SEQ ID: 1)

The 358 vector adopts hairpin conformations. Underlined bases indicate ribonucleic acid residues. The bolded letter indicates the mutator (heterologous region). The italicized "T"'s indicate the unpaired bases.

A vector having the same sequence but containing no ribonucleic acids was also constructed for use as a control. In the vector thymidine was substituted for uracil.

The control vector for mutation at base 358: (DNA 358 Vector) 5' TGCCGATCGGCAACTTTTGTGCGATCGGCAAATTT 3' (SEQ ID: 2) also adopts a hairpin conformation. Again, the bolded residue is the mutator.

An *in vivo* transformation experiment was carried out in which *U. maydis* cells ( $10^7$ ) were made into protoplasts with a recovery of  $10^6$  cells in a 50  $\mu$ l transformation buffer solution following methods well known in the art and transfected with different amounts of the chimeric vectors or homogeneous (DNA alone) vector were mixed with the protoplasts at 37°C. The cells were then plated on selective media and the number of surviving colonies (transformants)

were counted. The results are presented in tabular form below:

5	Transfection Amount ( $\mu$ g)	Vector - 358 Type	Number of Surviving Colonies
	0.1	DNA	0
	0.1	chimeric	13
	0.25	DNA	1
10	0.25	chimeric	49
	0.75	DNA	12
	0.75	chimeric	131
	1.0	DNA	19
	1.0	chimeric	573

15 These data show that transfection of cells with the RNA 358 Vector increases the survival of the cells at a much greater rate than transfection with the corresponding DNA vector.

## 20 6.2. Example 2: The Transformation of NIH 3T3 Cells.

NIH 3T3 cells are human cells that have a benign (controlled) growth characteristics. Malignant (uncontrolled) growth characteristics are conferred by the 25 single point mutation in the oncogene *H-ras* that replaces Gly<sup>12</sup> by Thr. Thus, the mutation G→T<sup>12</sup> leads to a readily selectable alteration in the growth characteristics. Taparowsky, E., et al., 1982, Nature 300:762; Sukumar S., et al., 1983, Nature 306:658.

30 Chimeric vectors were constructed to direct the G→T<sup>12</sup> mutation having the following sequence:  
5'-CACACCGACGGCGCCACCACTTTGTGGTGGGCGCCGTCGGTGTGGGTTTGCC-3'  
(SEQ ID: 3)

35 The sequence is presented in the conventional single letter code with the additional features the RNA (underlined), unpaired bases are italicized, and the mutator base is bold.



Note that after circularization the chimeric vector is divided into two substantially complementary strands by the two trithymidinyl sequences and that all the ribose containing nucleotides are present in only one of the two 5 strands.

Two forms of the chimeric vector were synthesized using 2'-OMe ribose bases having respectively one ("R") and two ("R-D-R") regions of hybrid duplex flanking four deoxyribose residues. The R-D-R form is shown above, SEQ ID: 10 3, the R form is identical except that bases 6-9 ("CGAC") were deoxynucleotides. Note that the 5' and 3' termini are deoxynucleotides. This allows the chimeric vectors to be circularized after synthesis by use of the same DNA ligase procedures as are commonly employed in recombinant DNA. 15 After ligation, the circularization chimeric vectors were isolated from the substrate by two iterations of preparative electrophoresis in D600 gel (AT Biochem, Malvern, PA).

Control vectors were: 1) the same sequence lacking hybrid-duplex (data shown, "homo-duplex"); 2) the unpaired 20 DNA (data shown "sDNA") having the sequence 5'-GCCCACACCGACGGCGCCACCAC-3' (SEQ ID 4); 3) the chimeric vector having no heterologous region and hence no mutator nucleotide (data not shown).

The experiment was conducted by transforming NIH 3T3 25 cells ( $1 \times 10^6$  cells) using an electroporation procedure. After electroporation, the cells were plated at a seeding density of  $4 \times 10^3$  cells/cm<sup>2</sup> and allowed to grow for 14 days in culture. Transformants were visualized by staining foci-forming cells with crystal violet. As a positive control, 30 the plasmid pT24 that encodes the T<sup>12</sup> form of H ras was employed. This control was used to determine the level of illegitimate recombination in the transfected NIH 3T3 cells. The experiment was repeated five times and a summary of the results are presented in tabular form below. In addition to 35 the results presented below the control experiments that used a chimeric vector having no mutator codon showed no transformed foci of NIH 3T3 cells.

	<u>Vector Type</u>	<u>Amt Transfected</u> (per 10 <sup>6</sup> cells)	<u>Transformants</u> <u>per 10<sup>6</sup> cells</u> (14 days)	
5	pT24 (positive control)	10 $\mu$ g	57	
	sDNA	1 $\mu$ g	2	
	sDNA	10 $\mu$ g	13	
	homo-duplex	1 $\mu$ g	0	
10	homo-duplex	10 $\mu$ g	2	
	chimeric	50ng	R-D-R 19	R 12
	chimeric	200ng	55	43
	chimeric	1 $\mu$ g	139	103

15        These results show that the chimeric vectors introduced a specific mutation by homologous recombination so as to transform mammalian cells. The rate of transformation in this experiment was greater than the rate of transformation by illegitimate recombination that was observed by

20 ~~transfection with pT24 positive control vector, which~~ contained an entire mutated *H-ras* gene. Thus, by use of chimeric vectors a rate of homologous recombination in a mammalian cell was achieved that was greater than the rate of illegitimate recombination.

25

30

35

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

5 (i) APPLICANT: Kmiec, Eric

(ii) TITLE OF INVENTION: Compounds and Methods for Site Directed  
Mutations in Eukaryotic Cells

10 (iii) NUMBER OF SEQUENCES: 4

## (iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Pennie & Edmonds  
(B) STREET: 1155 Avenue of the Americas  
(C) CITY: New York  
(D) STATE: NY  
15 (F) ZIP: 10036-2711

## (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk ,  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
20 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

## (vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US  
(B) FILING DATE:  
25 (C) CLASSIFICATION:

## (viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Friebel, Thomas E.  
(B) REGISTRATION NUMBER: 29,258  
(C) REFERENCE/DOCKET NUMBER: 7991-009

## 30 (ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (212) 790-9090  
(B) TELEFAX: (212) 869-8864/9741

## (2) INFORMATION FOR SEQ ID NO:1:

35

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 37 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: self complementary
- (D) TOPOLOGY: linear

5

(ii) MOLECULE TYPE: DNA/RNA

(ix) FEATURE:

- (A) NAME/KEY: -
- (B) LOCATION: 19..32
- (D) OTHER INFORMATION: /label= a  
/note= "RNA"

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

15

TGCCGATCGG CAACTTTTGU UGCCGAUCGG CAAATTT

37

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: self complementary
- (D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: DNA

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

TGCCGATCGG CAACTTTTGT TGCCGATCGG CAAATTT

37

(2) INFORMATION FOR SEQ ID NO:3:

30

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 53 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: self complementary
- (D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: DNA/RNA

## (ix) FEATURE:

- (A) NAME/KEY: -
- (B) LOCATION: 2..5
- (D) OTHER INFORMATION: /label= b  
/note= "RNA"

5

## (ix) FEATURE:

- (A) NAME/KEY: -
- (B) LOCATION: 10..21
- (D) OTHER INFORMATION: /label= d  
/note= "RNA"

10

## (ix) FEATURE:

- (A) NAME/KEY: -
- (B) LOCATION: 51..52
- (D) OTHER INFORMATION: /label= f  
/note= "RNA"

15

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CACACCGACG GCGCCACCA CTTGTGGTG GCGCCGTCG GTGTGGGTTT GCC

53

## 20 (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: self complementary
- (D) TOPOLOGY: linear

25

## (ii) MOLECULE TYPE: DNA

## 30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GCCCACACCG ACGGCGCCAC CAC

23

35

## Claims:

1. A chimeric nucleic acid having at most one 3' end and one 5' end, which nucleic acid further comprises:

5

- a. a region of contiguous unpaired bases disposed so that the unpaired region separates the nucleic acid into a first strand and a second strand;
- b. a region of Watson-Crick paired nucleic acid of at least 15 base pairs in length, in which bases of the first strand correspond to bases of the second strand, and in which;
- c. the first strand comprises a region of at least three contiguous nucleotides comprised of a 2'-O or 2'-OMe ribose, which are Watson-Crick paired in a hybrid-duplex.

10

15

2. The nucleic acid of Claim 1 in which the second strand contains no nucleotides comprised of a 2'-O or 2'-OMe ribose.

20

3. The nucleic acid of Claim 2 in which the hybrid-duplex comprises at least 9 contiguous nucleotides comprised of a 2'-O ribose flanked by nucleotides comprised of 2'-deoxyribose.

25

4. The nucleic acid of Claim 2 in which the first strand comprises two regions of contiguous nucleotides comprised of a 2'-O or 2'-OMe ribose that form two regions of hybrid-duplex, each hybrid-duplex region having at least 4 base pairs of length, and an interposed region of at least 1 base pair of homo-duplex disposed between the hybrid duplex regions.

30

5. The nucleic acid of Claim 2 in which the first strand comprises two regions of nucleotides comprised of a 2'-O or 2'-OMe ribose that form two regions of hybrid-duplex,

35

each hybrid-duplex region having at least 8 base pairs of length, and an interposed region of at least 4 base pairs of homo-duplex disposed between the hybrid duplex regions.

5           6.    The nucleic acid of Claim 5 in which the interposed region of homo-duplex is consists of between 4 and 50 2'-deoxyribose base pairs.

7.    The nucleic acid of Claim 5 in which the interposed  
10 region of homo-duplex is consists of between 30 and 1,000 2'-deoxyribose base pairs.

8.    The nucleic acid of Claim 2 in which the 5' and 3' ends are ligated.

15

9.    The nucleic acid of Claim 5 in which the 5' and 3' ends are ligated.

10.   The nucleic acid of Claim 6 in which the 5' and 3'  
20 ends are ligated.

11.   The nucleic acid of Claim 7 in which the 5' and 3' ends are ligated.

25       12.   A method of introducing a predetermined alteration in a target sequence of the genome of a cell, having a nucleus, which comprises the steps of:

- a.    providing a chimeric vector having two regions  
homologous with a target sequence and a  
30 heterologous region, disposed there between,  
encoding the alteration;
- b.    maintaining the chimeric vector within the nucleus  
of the cell such that the chimeric vector and the  
target sequence genetically recombine.

35

13. The method of claim 12 wherein the nucleus is an pronucleus of an ova and the introduction is by direct injection.

5 14. The method of claim 12 wherein the nucleus is a nucleus of an embryonic stem cell and the introduction is by direct injection.

15 15. A method of obtaining a cell population having altered characteristics which comprises the steps of:

a. providing a chimeric vector having two regions homologous with a target sequence and a heterologous region, disposed there between, encoding the alteration;

15 b. maintaining the chimeric vector within the nucleus of the cell whereby the chimeric vector and the target sequence recombine and the cell is transformed; and

c. selecting the transformed cells.

20

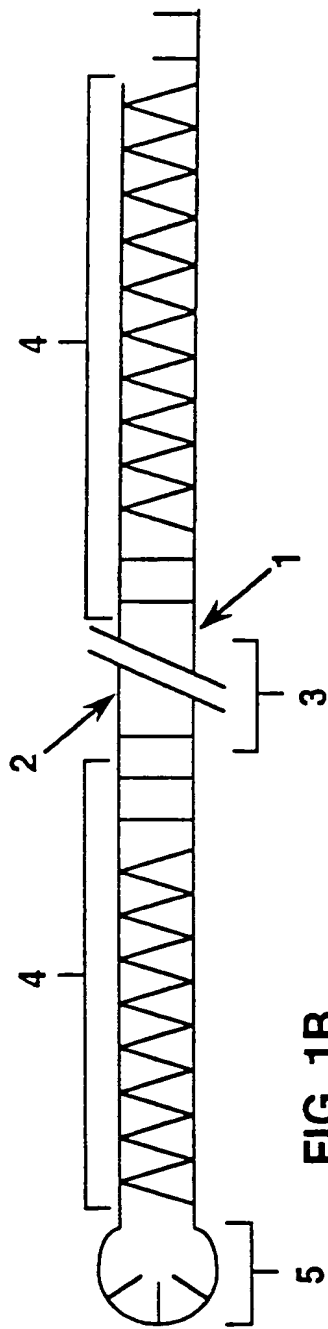
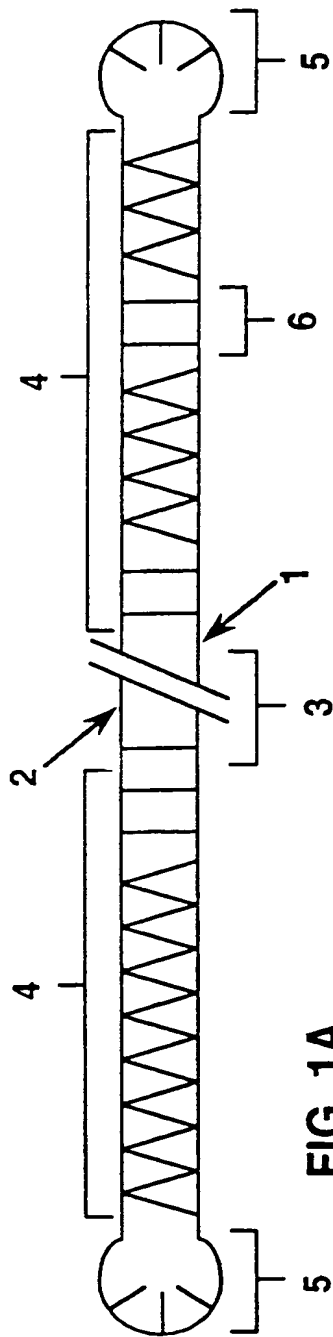
16. The method of claim 15 wherein the cell is a mammalian cell.

17. The method of claim 15 wherein the cell is a yeast  
25 or fungus.

30

35





## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US94/14181

### A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : C07H 21/02, 21/04; C12N 15/00

US CL :435/172.3, 320.1; 536/22.1

According to International Patent Classification (IPC) or to both national classification and IPC

### B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/172.3, 320.1; 536/22.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

## APS, DIALOG

### C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	Nature, Volume 344, issued 08 March 1990, Zheng et al., "Gene Targeting in Normal and Amplified Cell Lines", pages 170-173, see entire document.	1-17

☐ Further documents are listed in the continuation of Box C.

 See patent family annex.

•	Special categories of cited documents:	T	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*A*	document defining the general state of the art which is not considered to be of particular relevance		
*E*	earlier document published on or after the international filing date	*X*	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
*L*	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*Y*	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
*O*	document referring to an oral disclosure, use, exhibition or other means	*Z*	document member of the same patent family
*P*	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search

21 FEBRUARY 1995

Date of mailing of the international search report

**27 FEB 1995**

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Authorized officer

DOUGLAS GURLAN-SHERMAN

Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)\*